AD_____

Award Number: DAMD17-02-1-0246

TITLE: A Large Scale Genomic Approach to Prostate Cancer Risk in African-American Men

PRINCIPAL INVESTIGATOR: Matthew L. Freedman, M.D.

CONTRACTING ORGANIZATION: Massachusetts General Hospital
Boston, MA 02114

REPORT DATE: June 2005

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-06-2005	Annual Summary	1 Jun 02 – 31 May 05

4. TITLE AND SUBTITLE
A Large Scale Genomic Approach to Prostate Cancer Risk in African-American Men

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-02-1-0246

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Matthew L. Freedman, M.D.

E-Mail: mfreedman@partners.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts General Hospital
Boston, MA 02114

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Attached at next page.

15. SUBJECT TERMS
Genomics, Genetics, Prostate Cancer Risk, Admixture Mapping

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	11	19b. TELEPHONE NUMBER *(include area code)*

ABSTRACT

We are studying how genetic variation impacts sporadic prostate cancer risk in a multiethnic cohort with a particular emphasis on African American men. During the funding period, we have performed candidate gene and genome wide studies. Our candidate gene studies have focused on genes with a high likelihood of influencing prostate cancer and have been some of the largest and most thorough studies yet performed. We have demonstrated that: a) the androgen receptor, insulin like growth factor binding proteins 1 and 3 do not demonstrate a significantly increased risk of prostate cancer, b) variation at the insulin like growth factor 1 locus does contribute to increased risk. We have also performed the first genome wide admixture scan. We demonstrated that the average proportion of African ancestry is slightly, but significantly higher in cases than controls. We were unable, however, to conclusively identify a locus that accounted for the excess risk of prostate cancer in African American men. Our comprehensive and thorough approaches set the stage for better understanding how our genetic heritage influences prostate cancer risk.

Table of Contents

Cover.....

SF 298.....

Introduction.....4

Body.....5-8

Key Research Accomplishments.....9

Reportable Outcomes.....9-10

Conclusions.....10

Appendices.....

INTRODUCTION:

Prostate cancer is the most common non-cutaneous cancer in North American men. Moreover, African-American men suffer from the highest measured incidence of prostate cancer in the world. This increased risk is, in part, due to genetic factors. Despite having one of the highest heritabilities (genetic components) of all epithelial cancers, convincing evidence of loci involved in prostate cancer risk has been hard to come by. Furthermore, most studies are performed in largely White populations. We proposed to perform large scale genomic studies in a multiethnic population, including a substantial sample size of African-Americans. During the funding period, we examined population stratification, a potential confounder in association studies, and performed association studies using both a candidate gene approach and, more recently, a whole genome admixture mapping approach.

BODY:

Task 1. Evaluation of stratification in case and control populations in order to assess its impact on association studies.

We have exceeded the number of markers and evaluation of stratification as outlined in Task 1 in the Statement of Work in order to thoroughly study this topic. We intensively studied the African American population based on the hypothesis that stratification would be more likely to occur in this group than other populations since the prevalence of prostate cancer is higher in Africans than Europeans (although our study also studied other populations as well). Hence, African-American cases would be expected to possess a higher proportion of African ancestry than controls, leading to systematic differences in allele frequencies. Initially, we genotyped 46 markers in 93 cases and 86 controls to assess for stratification (Table 1). A summary χ^2 statistic revealed no significant stratification [5]. However, using a more quantitative metric, termed genomic control (GC), the data were still consistent with stratification [6]. This could be due to one of two scenarios: a) cryptic stratification is present (subtle degrees of stratification that are not adequately captured by self-reported ethnicity), or b) stratification is not present.

To differentiate between these alternatives, we increased our power to detect stratification by genotyping 210 markers in our sample of 469 African-American prostate cancer cases and 268 controls. We discovered that statistically significant stratification was, in fact, present in this study. Notably, this effect is present in a case-cohort designed study, which should be less susceptible to the effects of stratification. Although the magnitude of this effect may seem modest, it is expected to impact the false positive rate of a study, especially when trying to identify genetic variants that confer risk in a complex disease such as cancer. For example, with the estimated 95% upper bound on λ of 3.34 (what we actually measured in our sample) in a study of 1,000 cases / controls, we would expect that if 100,000 SNPs were tested, i.e., close to a whole genome scan, 1,568 false positive results would be expected due to stratification.

Our analysis reveals that population stratification affects case-control studies in practice, and that despite the uttermost care in matching cases and controls, it is likely to become increasingly important factor in case-control studies of the future, as sample sizes increase in order to detect more subtle genetic effects and correct for multiple hypothesis testing. We were also able to provide rational guidelines for how many SNPs would be needed to control for this source of confounding. The number of SNPs necessary to measure and correct for stratification depends critically on the magnitude of effect one observes at a candidate locus.

This study resulted in a manuscript that was recently published in *Nature Genetics* and is one of the largest empirical studies of population stratification to date. We definitively demonstrate that modest, but important levels of stratification are present and cannot be safely ignored. We further provide a methodology for measuring and controlling for the amount present in the sample.

(update 2005) Since this report has been published, a number of other studies have, in fact, confirmed our finding that non-negligible magnitudes of stratification exist.

Tasks 2 and 3: Obtain genetic variation information, in the form of SNPs, for each of the 39 genes to be investigated in the growth hormone pathway and evaluate marker assays. Assess genetic variation in genes in growth hormonal pathway.

Since the initial writing of this grant, genotyping costs have decreased and throughput has increased resulting in the consideration of ever increasing scales of association studies. We have transitioned from performing candidate gene studies to implementing a technique termed admixture mapping to scan the genome for prostate cancer susceptibility loci.

One of our first candidate gene projects involved systematically evaluating inherited variation at the androgen receptor (*AR*) locus and its relationship to prostate cancer risk in a multiethnic cohort. Numerous lines of experimental and epidemiologic evidence point to the *AR* as a plausible candidate gene for prostate cancer. Many studies report a positive association, while others report no association. Moreover, only one polymorphism (the CAG repeat in exon 1) had been extensively studied. Our goal was to identify and address the reasons for the inconsistent findings in the literature.

Our study was the largest and most thorough study of this locus. We genotyped the CAG microsatellite polymorphism in a set of 4,196 cases and controls (the largest and most comprehensive study to date) and found no significant association when analyzing this repeat as a continuous variable or as a cutpoint variable (see attached manuscript). We resequenced the exons in 88 advanced cases of prostate cancer and did not find any amino acid altering variants. To survey the noncoding region, we genotyped a total of 32 polymorphic SNPs across ~275 kb in a multiethnic population (African-American, Caucasian, Japanese and Latino). As noted in prior studies, the African-American population possesses the greatest diversity, i.e., 30 polymorphic markers. In sharp contrast, the Japanese population is monomorphic at all 32 sites. Thus, while the AA population has 14 haplotypes across this region, only 1 haplotype is segregating in the Japanese population. We tested these haplotypes in a large prostate cancer cohort (African-American, N=1,003, Caucasian, N=209, Japanese, N=242, and Latino, N=302). The haplotypes did not reveal any evidence of association with prostate cancer risk. Our results provide conclusively demonstrate that the AR does not play a significant role in prostate cancer predisposition. These findings were recently published in the *American Journal of Human Genetics*. Furthermore, these data are being used to inform the NCI funded Breast and Prostate Cancer and Hormone-Related Gene Variants Cohort Consortium, an initiative that pools samples from across 6 different cohorts yielding close to 9,000 prostate cancer cases.

Of note, we discovered some of the most extreme signatures of evolutionary selection in the vicinity of the *AR* as a result of this work. Genetic evidence of selection illuminates loci that were important in our species' history and may serve as loci that are involved in traits in the current day population. Results from this study have been submitted to *PLoS Genetics*.

The next candidate gene we assessed was the insulin like growth factor I (*IGFI*) gene. There is strong epidemiologic evidence that higher serum levels of IGFI are associated with prostate cancer risk; however, large scale studies of this locus have not been undertaken. We genotyped 64 SNPs across approximately 155 kb (~69kb 5' of the locus, 63 kb - the IGF1 locus, and ~23 kb 3' of the locus) (figure 1). Twenty-nine (out of the 64) SNPs were needed to

accurately "tag" the variation at this locus. A range of 5-12 common (>.05 in any one ethnic group) haplotypes were observed in the population.

Our results demonstrate an association between inherited variation at IGFI and prostate cancer risk. Our analysis points to a variant (rs7965399) that can explain most of the risk P=0.002. Risk did not significantly vary across ethnic groups. As the association literature is replete with false positive results (at least partially due to overly liberal thresholds of significance, i.e., P=0.05), we also provide a framework for evaluating the likelihood that a particular variant is truly involved in conferring an elevated risk. Permutation testing is performed as well as criteria set forth in a study that incorporates a method for multiple hypothesis testing (J Natl Cancer Inst. 2004 Mar 17;96(6):434-42). These findings are in press at the *Journal of the National Cancer Institute*.

We have currently completed the haplotype structure in PIK3CG, IGFBP1, IGFBP3, and IGF2R. The SNP databases have grown considerably over the past year making us more confident that we have captured the majority of variation at a locus. A report on IGFBP1 and IGFBP3 (negative results) has been written and will be submitted for publication in the next month.

As noted above, during the course of this grant, genotyping costs have decreased and throughput has increased to the point where even larger scale studies can be performed. Ideally, genome wide studies are ideal because they do not require any *a priori* assumptions and instead of studying one gene at a time, the whole genome can be tested. Comprehensive scans at the resolution of the candidate genes highlighted above are still too expensive. A technique termed admixture mapping, however, *can* scan the whole genome with 2-3 orders of magnitude less markers (and therefore less cost). Admixture, as its name implies, is a technique that requires a population (e.g., African American) derived from two or more previously isolated parental populations (e.g., West African and European). Although the concept of admixture mapping has been known for decades, the tools (methodology and polymorphisms with particular characteristics) have only recently become available that are required to utilize this technique. Of note, admixture mapping is only powered to detect variants that differ in frequency between the parental populations. Therefore, diseases such as prostate cancer in African American men (an admixed population) where the prevalence of the disease is higher presumably due, at least in part, to genetics are particularly strong candidates for admixture mapping.

In close collaboration with David Reich, a population geneticist at Harvard, we performed the first genome wide admixture mapping scan in sporadic prostate cancer in an African American population. We genotyped over 1,000 markers in over 2,035 samples. Our results demonstrate a modest, but significant increase in the average proportion of African ancestry in cases versus controls. This is one of the first demonstrations that the increased prevalence of prostate cancer is in fact due to risk variants that are more frequent in the African than European population. While we did not discover any single locus that achieved our pre-determined level of genome wide significance, we were able to rule out every locus in the genome as conferring more than 1.5-fold increased risk due to African ancestry. We conclude that the increased genetic risk for prostate cancer in African Americans must be spread over multiple loci, which each individually confer too little risk to be detectable with the current sample size. Currently, we are genotyping more samples and are in the process of preparing a manuscript describing our findings.

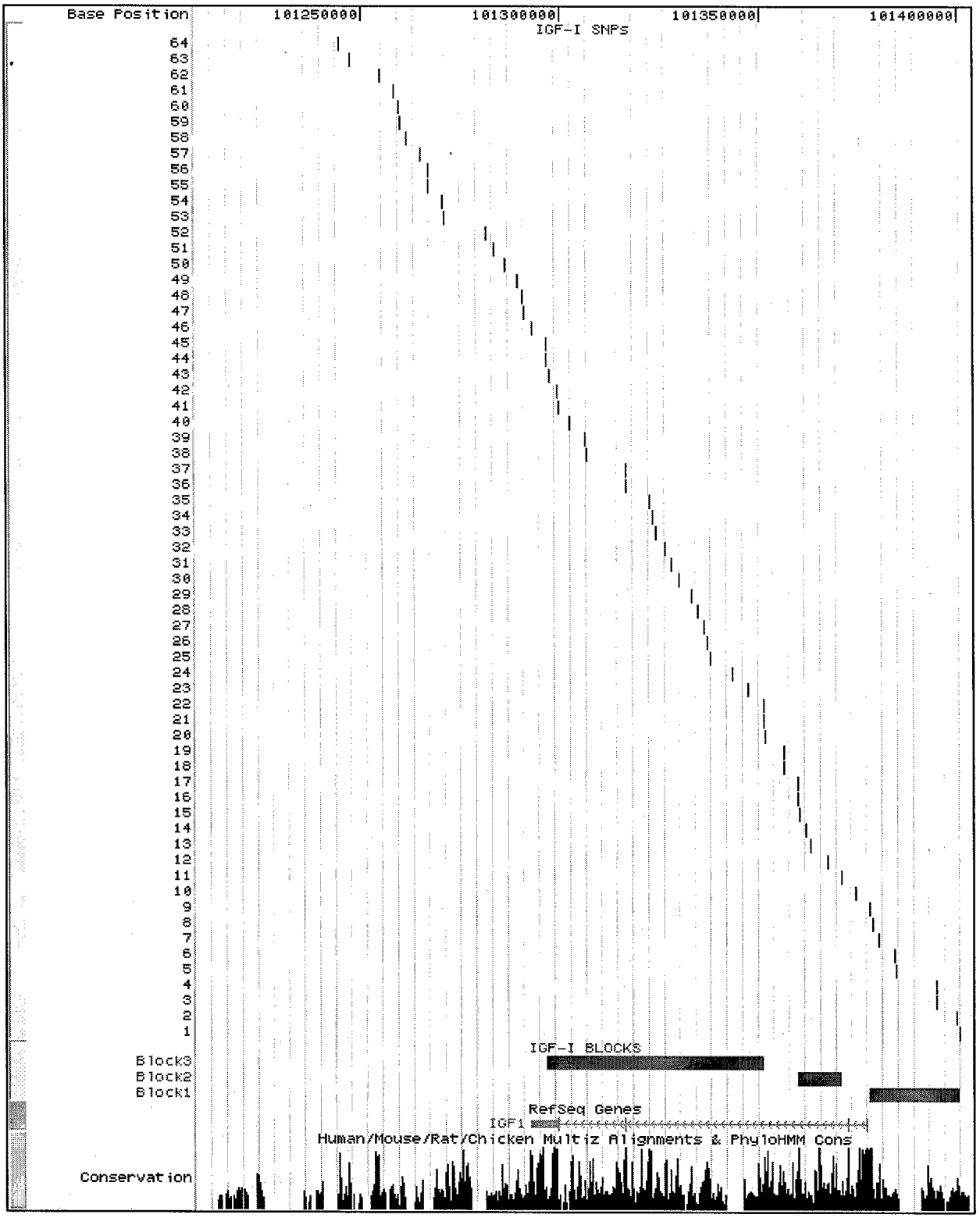

Figure 1: UCSC browser (http://genome.ucsc.edu) snapshot of position of snps and blocks relative to the IGF-1 locus.

KEY RESEARCH ACCOMPLISHMENTS:

1. Comprehensive survey of genetic variation in the androgen receptor gene demonstrating lack of strong association to prostate cancer settling decades of research. Findings published in the *American Journal of Human Genetics*.

2. Research directly stemming from Accomplishment number 1 demonstrates the area of the genome adjacent to the androgen receptor was under recent intense selection, highlighting this as an important location in our genome. Manuscript submitted to *PLoS Genetics*.

3. Demonstration of a positive association in inherited variation at the *IGFI* locus and prostate cancer risk. Manuscript in press at the *Journal of the National Cancer Institute*.

4. Thorough characterization and testing of *IGFBP1* and *IGFBP3* loci in a large multiethnic cohort. Genetic variation at these loci is not significantly associated with prostate cancer risk.

5. Performed the first genome wide admixture scan for sporadic prostate cancer in over 2,000 African American men. Our results convincingly demonstrate that African American men with prostate cancer have a small, but significant increase in African ancestry relative to controls. We did not discover any locus conferring a genome wide significant increase in prostate cancer risk using this novel method. We conclude that more samples will be necessary to detect these risk variants.

REPORTABLE OUTCOMES:

1. Abstract presentation on IGF1 at 2005 American Association for Cancer Research (AACR) Annual Meeting

2. Invited speaker at the 2005 AACR annual meeting - "The impact of population stratification on genetic association studies"

3. Invited speaker at the 4th Era of Hope meeting for the Department of Defense (DOD) Breast Cancer Research Program (BCRP) – "Large scale genomic association studies in breast cancer"

4. Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. *Am J Hum Genet.* 2005 Jan;76(1):82-90.

5. Association study of common genetic variation in *IGF1* and prostate cancer risk in the Multiethnic Cohort, *manuscript in press at the Journal of the National Cancer Institute.*

CONCLUSIONS:

1. The androgen receptor data reveals striking differences in haplotype frequencies between different populations. These observations are extremely important, especially when studying diseases such as prostate cancer that demonstrate clear ethnic predispositions. We have generated the largest and most comprehensive study to date on common variation in the androgen receptor and prostate cancer risk and demonstrate no demonstrative association.

2. We find a significant association between genetic variation in *IGF1* and prostate cancer risk.

3. We do not find significant evidence for association between *IGFBP1* nor *IGFBP3* and prostate cancer risk.

4. The average proportion of African ancestry is modestly, but significantly increased in African American men with prostate cancer relative to African American controls.

5. In the first genome wide admixture scan, we did not discover any single locus that achieved our pre-determined level of genome wide significance. Our well-powered study, however, ruled out every locus in the genome as conferring more than 1.5-fold increased risk due to African ancestry thus providing important guidelines for future studies.